

Low Dimensional Atomic Norm Representations in Line Spectral Estimation

Maxime Ferreira Da Costa and Wei Dai

Department of Electrical and Electronic Engineering, Imperial College London, United Kingdom

Email: {maxime.ferreira, wei.dai1}@imperial.ac.uk

Abstract—The line spectral estimation problem consists in recovering the frequencies of a complex valued time signal that is assumed to be sparse in the spectral domain from its discrete observations. Unlike the gridding required by the classical compressed sensing framework, line spectral estimation reconstructs signals whose spectral supports lie *continuously* in the Fourier domain. If recent advances have shown that atomic norm relaxation produces highly robust estimates in this context, the computational cost of this approach remains, however, the major flaw for its application to practical systems.

In this work, we aim to bridge the complexity issue by studying the atomic norm minimization problem from low dimensional projection of the signal samples. We derive conditions on the sub-sampling matrix under which the partial atomic norm can be expressed by a low-dimensional semidefinite program. Moreover, we illustrate the tightness of this relaxation by showing that it is possible to recover the original signal in *poly-logarithmic time* for two specific sub-sampling patterns.

I. INTRODUCTION

A. Background on line spectral estimation

The line spectral estimation problem aims to recover the frequencies of a complex time signal that is assumed to be sparse in the spectral domain from its discrete measurements $x \in \mathbb{C}^n$, uniformly acquired at a sampling frequency F_s . More precisely, we suppose that the sampled time signal is supported over a small number of frequencies s^* and that we dispose of observations $x \in \mathbb{C}^n$ of the form

$$\forall j \in \llbracket 0, n-1 \rrbracket, \quad x_j = \sum_{k=1}^{s^*} c_k^* e^{i2\pi f_k^* j}, \quad (1)$$

whereby $\{f_k^*\}_{1 \leq k \leq s^*}$ is the ordered set containing the s^* spectral components generating the signal x , and $\{c_k^*\}_{1 \leq k \leq s^*}$ the one of their associated complex amplitudes. Both of those sets, as well as their cardinality s^* , are supposed to be unknown. We highlight at this point that the particularly of this model is that the frequencies $\{f_k^*\}$ can be drawn *continuously* in $[0, 1)$ and *are not constrained to belong to some finite discrete grid*, unlike in the classic compressed sensing framework. The ground truth spectral distribution of x , denoted \hat{x}^* , is therefore constituted of s^* Dirac spikes located at the frequencies $\{f_k^*\}$ forming the spectral support of \hat{x}^* .

It is not difficult to see that this problem is ill-posed, in the sense that there are infinitely many estimators of the ground truth spectral distribution \hat{x}^* that are consistent with the measurement vector x . Among all of those estimators,

the one considered to be optimal in the line spectral estimation framework is the one returning the consistent spectral distribution \hat{x}_0 having the sparsest possible spectral support. Alternatively, this estimator can be formulated as the output of the *non-convex* minimization program

$$\hat{x}_0 = \arg \min_{\hat{x} \in D_1} \|\hat{x}\|_0, \quad \text{subject to } x = \mathcal{F}_n(\hat{x}), \quad (2)$$

whereby D_1 denotes the set of tempered integrable spectral distributions, $\mathcal{F}_n : D_1 \rightarrow \mathbb{C}^n$ is the inverse discrete-time Fourier operator, and $\|\cdot\|_0$ is the form mapping D_1 to $[0, \infty]$ counting the (potentially infinite) cardinality of the spectral support.

The study of line spectral recovery under the paradigm of convex relaxations has been gaining in popularity after the tightness results of such approaches were demonstrated in the pioneer works [1], [2], [3]. It has been shown in [4] that a separation criterion on the spectral support of \hat{x}^* of the form $\Delta_{\mathbb{T}}(\hat{x}^*) = \min_{i \neq j} \{|f_i^* - f_j^*|\} \geq \frac{C_{\text{nec}}}{n-1}$ is always necessary to guarantee the tightness of convex approaches for a constant $C_{\text{nec}} = \frac{1}{\pi}$. On the other hand, it was shown in [1] and enhanced in [5] that $C_{\text{suf}} = 2.56$ is sufficient to ensure the recoverability of \hat{x}_0 from a convex surrogate. Since then, many direct extensions of Program (2) were proposed. Among them, we mention an extension to multi-dimensional spectra [6], [7], to multiple measurement vectors [8], and to the spectral blind deconvolution framework [9]. Line spectrum estimation theory can be viewed as a particular case of the spikes deconvolution problem presented in [10].

B. Notations

The adjunction of \mathbf{X} is denoted \mathbf{X}^* , whenever \mathbf{X} is a vector, a matrix, or a linear operator. The transposition of \mathbf{X} is written \mathbf{X}^T . The set of complex square matrices of dimension n is denoted $M_n(\mathbb{C})$ and vectors of \mathbb{C}^n are indexed in $\llbracket 0, n-1 \rrbracket$ so that every vector $u \in \mathbb{C}^n$ writes $u = [u_0, \dots, u_{n-1}]^T$. The trace operator is denoted $\text{tr}(\cdot)$. We define by $\mathcal{T}_n : \mathbb{C}^n \rightarrow M_n(\mathbb{C})$ the Hermitian Toeplitz generator in dimension n , such that for all $u \in \mathbb{C}^n$, $\mathcal{T}_n(u)$ is the Hermitian Toeplitz matrix whose first row is equal to u . Its adjoint \mathcal{T}_n^* is characterized for every matrix $H \in M_n(\mathbb{C})$ by

$$\forall k \in \llbracket 0, n-1 \rrbracket, \quad \mathcal{T}_n^*(H)[k] = \langle \Theta_k, H \rangle = \text{tr}(\Theta_k^* H),$$

whereby Θ_k is the elementary Toeplitz matrix equals to 1 on the k^{th} upper diagonal and zero elsewhere. Moreover, for every

matrix $M \in \mathbb{C}^{m \times n}$, $m \leq n$, we denote by \mathcal{R}_M the operator given by

$$\begin{aligned} \mathcal{R}_M : \mathbb{C}^m &\rightarrow \mathbb{M}_m(\mathbb{C}) \\ v &\mapsto \mathcal{R}_M(v) = M\mathcal{T}_n(M^*v)M^*. \end{aligned}$$

Its adjoint \mathcal{R}_M^* is consequently characterized for every matrix $S \in \mathbb{M}_m(\mathbb{C})$ by $\mathcal{R}_M^*(S) = M\mathcal{T}_n^*(M^*SM)$.

C. Atomic norm minimization

Atomic norms were analyzed in [11] as a generic way to regularize sparse inverse problems defined over continuous dictionaries. The underlying idea consists in considering the dictionary of interest \mathcal{A} as a set of building blocks called “atoms”, and to endow the search space E with the norm induced with the gauge of \mathcal{A} defined by

$$\forall x \in E, \|x\|_{\mathcal{A}} = \inf_{t > 0} \{x \in t \operatorname{conv}(\mathcal{A})\}. \quad (3)$$

The atomic ball defined by $\{x \in E : \|x\|_{\mathcal{A}} \leq 1\}$ is by construction the smallest convex set containing the dictionary \mathcal{A} , and one may expect, by analogy with ℓ_1 minimization in the discrete compressed sensing framework, that atomic norm minimization has a high sparsity promoting power.

Atomic norms have been introduced in the context of line spectral estimation in [3], [12]. Detailed performance guarantees of this use can be found in [13]. In this context, the set of underlying atoms for Model (1) takes the form $\mathcal{A} = \{a(f, \phi), f \in [0, 1), \phi \in [0, 2\pi)\}$, where each atom $a(f, \phi) \in \mathbb{C}^n$ writes for every $f \in [0, 1)$ and $\phi \in [0, 2\pi)$

$$\begin{aligned} a(f, \phi) &= \left[e^{i\phi}, e^{i(2\pi f + \phi)}, \dots, e^{i(2\pi(n-1)f + \phi)} \right]^T \\ &= e^{i\phi} a(f, 0). \end{aligned}$$

Using Carathéodory’s theorem on convex hulls, any vector of $\operatorname{conv}(\mathcal{A})$ can be expressed by a convex combination of at most $n + 1$ points in \mathcal{A} and the atomic norm reformulates

$$\|x\|_{\mathcal{A}} = \inf_{\substack{c_k \geq 0, \phi_k \in [0, 2\pi) \\ f_k \in [0, 1)}} \left\{ \sum_k c_k, x = \sum_k c_k a(f_k, \phi_k) \right\}. \quad (4)$$

Atomic norm regularization consists in approaching the pair of sets $(\{f_k^0\}, \{c_k^0\})$ characterizing the distribution \hat{x}_0 with the pair $(\{f_k^A\}, \{c_k^A e^{i2\pi\phi_k^A}\})$ minimizing (4) for a given observation vector x . It was shown in [12] that this estimate is exact, under proviso of a minimal spectral separation discussed earlier. Moreover, it was proven that one has the equality

$$\|x\|_{\mathcal{A}} = \inf_{\substack{u \in \mathbb{C}^n \\ t > 0}} \left\{ \frac{1}{2n} \operatorname{tr}(\mathcal{T}_n(u)) + \frac{1}{2}t : \begin{bmatrix} \mathcal{T}_n(u) & x \\ x^* & t \end{bmatrix} \succeq 0 \right\}, \quad (5)$$

where the infimum is reached over a pair $\{u_{\mathcal{A}}, t_{\mathcal{A}}\}$ for which the positive Toeplitz matrix $\mathcal{T}_n(u_{\mathcal{A}})$ admits for eigenvectors the atoms $\{a(f_k^A, 0)\}$. Hence, Relation (5) provides a direct way to recover the spectral support of \hat{x}_0 by solving a *convex* semidefinite program (SDP) of dimension $n + 1$.

II. PARTIAL LINE SPECTRAL ESTIMATION

A. Problem statement

We introduce the partial line spectral estimation problem by extension of the settings presented in Section I. The sampled vector $x \in \mathbb{C}^n$ generated according to Model (1) is now assumed to be unknown and one observe instead linear combinations $y \in \mathbb{C}^m$ of x through a known fat measurement matrix $M \in \mathbb{C}^{m \times n}$, so that $y = Mx$.

Identically to the original problem, our aim is to recover the sparsest spectral distribution $\hat{x}_{M,0} \in D_1$ matching the measurement vector y , which can be interpreted as the solution of the non-convex minimization program

$$\hat{x}_{M,0} = \arg \min_{\hat{x} \in D_1} \|\hat{x}\|_0, \text{ subject to } y = M\mathcal{F}_n(\hat{x}). \quad (6)$$

One could, of course, solve Problem (6) by regularizing it on \mathcal{A} , generalizing the approach followed in [3]; and add the extra linear measurement constraint $y = Mx$ to Relation (5), leading to the semidefinite program

$$\begin{aligned} (u_{M,\mathcal{A}}, t_{M,\mathcal{A}}) &= \arg \min_{\substack{u \in \mathbb{C}^n \\ t > 0}} \frac{1}{2n} \operatorname{tr}(\mathcal{T}_n(u)) + \frac{1}{2}t \\ \text{subject to } &\begin{bmatrix} \mathcal{T}_n(u) & x \\ x^* & t \end{bmatrix} \succeq 0 \\ &y = Mx. \end{aligned} \quad (7)$$

One can recover in a second time the spectral support via an eigen-decomposition of $\mathcal{T}_n(u_{M,\mathcal{A}})$. However, the SDP (7) involves a cost function and matrix constraint of dimension n , and such approach would require about $\mathcal{O}(n^7)$ operations using standard interior point methods such as SDPT3 or SUDEMI, whereas the *essential dimension* of the problem is of order m , potentially much smaller than n .

B. Contributions

The rest of this work aims to study two problematics emerging from the statement of the non-convex Program (6).

- How to *efficiently* solve Program (6) via atomic norm minimization? And, in particular, is there a semidefinite representation of dimension m for this problem?
- Can we find sub-sampling matrices M for which *exact recovery* of the spectral support of \hat{x}_0 is possible from the sole observation of the vector y ?

The first question is answered in Section III, where we derive, after proving a novel extension of the Carathéodory-Toeplitz lemma, conditions on the sub-sampling matrix M under which the partial line spectral estimation problem can be assimilated to an SDP of dimension $m + 1$. In Section IV, we provide theoretical tightness guarantees obtained by generalizing the Lagrange dual properties studied in [1], [10] onto the partial measurement case. At last, in Section V, we review certain sub-sampling patterns for which it is possible to guarantee the recoverability of the spectral support of \hat{x}^* from partial measurements. We illustrate that *poly-logarithmic time* recovery of \hat{x}^* is possible, and point out the *sub-Nyquist sampling* capabilities of such approaches.

III. MAIN RESULTS

A. Partial atomic norm relaxation

Given a sub-measurement matrix $M \in \mathbb{C}^{m \times n}$ with $m \leq n$, we define the associated partial atomic set \mathcal{B}_M as follows

$$\mathcal{B}_M = M\mathcal{A} = \{b_M(f, \phi), f \in [0, 1], \phi \in [0, 2\pi)\},$$

where $b_M(f_k, \phi_k) = Ma(f_k, \phi_k)$. Due to the absence of ambiguity on the matrix M in this work, the notations will be simplified to $\mathcal{B}_M = \mathcal{B}$ and $b_M(f_k, \phi_k) = b(f_k, \phi_k)$. Similarly to Equations (3) and (4), the partial atomic norm $\|\cdot\|_{\mathcal{B}}$ is defined for all $y \in \mathbb{C}^m$ by the gauge of \mathcal{B} , i.e.

$$\begin{aligned} \|y\|_{\mathcal{B}} &= \inf_{t>0} \{x \in t \operatorname{conv}(\mathcal{B})\} \\ &= \inf_{\substack{c_k \geq 0, \phi_k \in [0, 2\pi) \\ f_k \in [0, 1]}} \left\{ \sum_k c_k, y = \sum_k c_k b(f_k, \phi_k) \right\}. \end{aligned} \quad (8)$$

Once again the partial atomic ball verifies the property of being the smallest convex body containing all the building blocks $b(f, \phi)$. Moreover, the partial atomic norm relaxation consists in estimating the spectral support of $\hat{x}_{M,0}$ with the triplet of sets $(\{f_k^{\mathcal{B}}\}, \{c_k^{\mathcal{B}}\}, \{\phi_k^{\mathcal{B}}\})$ realizing the infimum of (8).

B. Semidefinite representability of partial atomic norms

The theoretical description (8) of partial atomic norm minimization does not provide an efficient way to compute an atomic decomposition on \mathcal{B} . On the other hand, the semidefinite representability of the atomic decomposition on \mathcal{A} provided in Equation (5) holds very specifically in the line spectral estimation framework due to the close link between complex exponential vectors of \mathbb{C}^n and Toeplitz matrices of same dimension. It was demonstrated by Carathéodory in [14] that the eigenvectors of positive Hermitian Toeplitz matrices are always elements of the form $a(f, 0)$.

In this section we propose in Lemma 1 a novel extension of Carathéodory's result, establishing a relationship between the elements of the form $b(f, 0) \in \mathcal{B}$ and the linear operator \mathcal{R}_M . We latter conclude in Theorem 2 on the semidefinite representability of the norm $\|\cdot\|_{\mathcal{B}}$.

Lemma 1 (Partial Carathéodory-Toeplitz lemma). *Consider a matrix $M \in \mathbb{C}^{m \times n}$, $m \leq n$, satisfying the following properties:*

- M is full rank.
- $\mathcal{T}_n^*(M^*M)$ belongs to the range of M^*

then any positive matrix S of rank r in the range of the operator \mathcal{R}_M can be decomposed under the form

$$S = VDV^*,$$

where

$$\begin{aligned} V &= [b(f_1, 0), \dots, b(f_r, 0)] \\ D &= \operatorname{diag}([d_1, \dots, d_r]), \end{aligned}$$

and whereby $d_k > 0$ are all positive real numbers.

Proof: Every vector $u \in \mathbb{C}^n$ can be decomposed under the form

$$u = w + z,$$

where $w \in \operatorname{range}(M^*)$ and $z \in \operatorname{range}(M^*)^\perp = \ker(M)$. Writing $w = M^*v$, it comes by linearity, for any $u \in \mathbb{C}^n$

$$\mathcal{T}(u) = \mathcal{T}(M^*v) + \mathcal{T}(z)$$

and

$$\begin{aligned} M\mathcal{T}(u)M^* &= M\mathcal{T}(M^*v)M^* + M\mathcal{T}(z)M^* \\ &= \mathcal{R}_M(v) + M\mathcal{T}(z)M^*. \end{aligned} \quad (9)$$

Moreover, using the properties of the matrix M , one has

$$\begin{aligned} M\mathcal{T}(z)M^* &= \langle M^*, \mathcal{T}(z)M^* \rangle \\ &= \langle \mathcal{T}^*(M^*M), z \rangle. \end{aligned} \quad (10)$$

Since $\mathcal{T}^*(M^*M)$ belongs by assumption to the range of the matrix M^* , its inner product with the vector z is null. Consequently, combining the previous (9) and (10), one gets

$$M\mathcal{T}(u)M^* = \mathcal{R}_M(v). \quad (11)$$

Now, let $S = \mathcal{R}_M(v)$ be a positive matrix in the range of \mathcal{R}_M . Using a dimension argument, there must exist at least one vector $z \in \operatorname{range}(M^*)^\perp$ for which the completion $\mathcal{T}(u) = \mathcal{T}(M^*v) + \mathcal{T}(z)$ is a positive matrix. By application of the Carathéodory-Toeplitz lemma [14], there exist a matrix $U = [a(f_1, 0), \dots, a(f_r, 0)]$ and a positive diagonal matrix $D = \operatorname{diag}([d_1, \dots, d_r])$ such that

$$\mathcal{T}(u) = UDU^*.$$

One concludes using Equation (11) that

$$\begin{aligned} \mathcal{R}_M(v) &= MU D U^* M^* \\ &= V D V^*, \end{aligned}$$

by letting $V = MU$, which concludes the proof. \blacksquare

We are now ready to state the main contribution of this work. The proof structure of this result is close to the one presented in [3, Proposition II.1] for the case of fully observed atomic norm minimization problems.

Theorem 2 (Semidefinite representability of partial atomic sets). *Suppose that $M \in \mathbb{C}^{m \times n}$ satisfies the conditions of Lemma 1 and that $\sqrt{n}M$ is a unitary matrix, then for any vector $y \in \mathbb{C}^m$, the following equality holds*

$$\|y\|_{\mathcal{B}} = \inf_{v \in \mathbb{C}^m} \left\{ \frac{1}{2m} \operatorname{tr}(\mathcal{R}_M) + \frac{1}{2}t, \begin{bmatrix} \mathcal{R}_M(v) & y \\ y^* & t \end{bmatrix} \succeq 0 \right\}. \quad (12)$$

Proof: First of all, since $\sqrt{n}M$ is unitary, one has $\|b(f, \phi)\|_2^2 = \operatorname{tr}(Ma(f, \phi)a(f, \phi)^*M^*) = \operatorname{tr}(I_m) = m$ for all $f \in [0, 1]$ and $\phi \in [0, 2\pi)$.

Denote by $\operatorname{SDP}(y)$ the quantity on the right hand side of (12). Suppose a decomposition of $y \in \mathbb{C}^m$ under the form $y = \sum_k c_k b(f_k, \phi_k)$ with $c_k > 0$, and denote by $v \in \mathbb{C}^m$

the phaseless counterpart of y given by $v = \sum_k c_k b(f_k, 0)$. Moreover, let $t = \sum_k c_k$ and $u = \sum_k c_k a(f_k, 0)$. Using (11)

$$\begin{aligned}\mathcal{R}_M(v) &= M \sum_k c_k a(f_k, 0) a(f_k, 0)^* M^* \\ &= M \sum_k c_k a(f_k, \phi_k) a(f_k, \phi_k)^* M^* \\ &= \sum_k c_k b(f_k, \phi_k) b(f_k, \phi_k)^*.\end{aligned}\quad (13)$$

Thus, the matrix form of interest can be identified as follows

$$\begin{bmatrix} \mathcal{R}_M(v) & y \\ y^* & t \end{bmatrix} = \sum_k c_k \begin{bmatrix} b(f_k, \phi_k) \\ 1 \end{bmatrix} \begin{bmatrix} b(f_k, \phi_k) \\ 1 \end{bmatrix}^*,$$

and is therefore a positive Hermitian matrix. Equation (13) ensures that $\text{tr}(\mathcal{R}_M(v)) = m \sum_k c_k$, and therefore that $\text{SDP}(y) \leq \sum_k c_k$. Consequently one has $\|y\|_{\mathcal{B}} \geq \text{SDP}(y)$.

It remains to show $\|y\|_{\mathcal{B}} \leq \text{SDP}(y)$ to conclude the proof. Suppose that there exist some $v \in \mathbb{C}^m$ such that

$$\begin{bmatrix} \mathcal{R}_M(v) & y \\ y^* & t \end{bmatrix} \succeq 0, \quad (14)$$

which implies that $\mathcal{R}_M(v) \succeq 0$. By application of Lemma 1, it is possible to decompose $\mathcal{R}_M(v)$ under the form

$$\mathcal{R}_M(v) = V D V^* = \sum_k d_k b(f_k, 0) b(f_k, 0)^*,$$

and the relationship $\text{tr}(\mathcal{R}_M(v)) = m \text{tr}(D)$ holds. Observing that (14) implies that y lies in the range of V , there exists a vector $w \in \mathbb{C}^m$ such that $y = \sum_k w_k b(f_k, 0) = V w$. Using the Schur complement lemma, it holds that

$$V D V^* \succeq \frac{1}{t} y y^* = \frac{1}{t} V w w^* V^*.$$

Call $q \in \mathbb{C}^m$ the vector solution of $V^* q = \text{sign}(w)$, which exists since V^* is full rank. We have that

$$\begin{aligned}\text{tr}(D) &= q^* V D V^* q \\ &\succeq \frac{1}{t} q^* V w w^* V^* q \\ &= \frac{1}{t} \left(\sum_k |w_k| \right)^2.\end{aligned}$$

This implies using the geometric mean comparison lemma that

$$\begin{aligned}\frac{1}{2m} \text{tr}(\mathcal{R}_M(v)) + \frac{1}{2} t &= \frac{1}{2} \text{tr}(D) + \frac{1}{2} t \\ &\geq \sqrt{\text{tr}(D) t} \\ &\geq \sum_k |w_k| \geq \|y\|_{\mathcal{B}},\end{aligned}$$

which completes the proof. \blacksquare

Remark 3. Lemma 1 and Theorem 2 both require the sub-sampling M to satisfy a bilinear relationship of the form $\mathcal{T}_n^*(M^* M) - M^* v = 0$ for some $v \in \mathbb{C}^m$. Although it is challenging to explicit the set of matrices satisfying this property, it is trivial to verify this hypothesis for a given matrix M . Moreover, many practical sub-sampling patterns do

satisfy this relation. For instance, consider a selection matrix $M = C_{\mathcal{I}}$ whose rows are equal to $\{e_j, j \in \mathcal{I}\}$ for some subset $\mathcal{I} \subseteq \llbracket 0, n-1 \rrbracket$ for cardinality m . This category of sub-sampling matrices corresponds to practical signal processing sampling schemes where the sample x_j is either kept, if $j \in \mathcal{I}$, either discarded. It can be easily verified that

$$\begin{aligned}\mathcal{T}_n^*(C_{\mathcal{I}}^* C_{\mathcal{I}}) &= \mathcal{T}_n^*(\text{diag}(\mathcal{I})) \\ &= m e_0,\end{aligned}$$

whereby $e_0 \in \mathbb{C}^n$ is the first vector of the canonical basis. Therefore, it comes that $C_{\mathcal{I}}$ satisfies the desired properties if and only if $0 \in \mathcal{I}$. Examples involving such sub-sampling matrices will be discussed in more details in Section V.

Finally, combining the definition of the partial atomic relaxation (8) with Theorem 2 leads to the following corollary.

Corollary 4. *If the sub-sampling matrix $M \in \mathbb{C}^{m \times n}$ satisfies the conditions of Theorem 2, then the estimate $\hat{x}_{\mathcal{B}}$ of the spectral support of \hat{x}_0 can be computed by solving the semidefinite program*

$$\begin{aligned}(v_{\mathcal{B}}, t_{\mathcal{B}}) &= \arg \min_{\substack{v \in \mathbb{C}^m \\ t > 0}} \frac{1}{2m} \text{tr}(\mathcal{R}_M(v)) + \frac{1}{2} t \\ \text{subject to} \quad &\begin{bmatrix} \mathcal{R}_M(v) & y \\ y^* & t \end{bmatrix} \succeq 0,\end{aligned}\quad (15)$$

whereby the estimated supporting frequencies $\{f_k^{\mathcal{B}}\}$ can be recovered from the vectors $\{b(f_k^{\mathcal{B}}, 0)\}$ forming the columns of $V \in \mathbb{C}^{m \times r}$ where $\mathcal{R}_M(v_{\mathcal{B}}) = V D V^*$. Moreover, the equality $M u_{\mathcal{A}, M} = v_{\mathcal{B}}$ holds between the outputs Program (7) and Program (15).

The semidefinite program (15) is of dimension $m+1$, and can be solved in polynomial time with respect to m using appropriate out of the box convex solvers or the Alternating Direction Method of Multipliers.

IV. TIGHTNESS OF PARTIAL ATOMIC RELAXATION

The guarantees provided in [1], [5], [10] for the full line spectral estimation problem are based on the existence of a polynomial obeying certain extremal properties. Such polynomial is often refereed as *dual certificate* for the Program (2). Its existence suffices to guarantee to tightness of the atomic relaxation as well as the uniqueness of the solution. The next proposition extends this theory to the partial observation case for an arbitrary matrix $M \in \mathbb{C}^{m \times n}$.

Proposition 5 (Dual certifiability). *If there exists a polynomial $Q_{\star} \in \mathbb{C}^{n-1}[X]$ having for coefficients vector $q_{\star} \in \mathbb{C}^n$ satisfying the conditions*

$$\begin{cases} q_{\star} \in \text{range}(M^*) \\ Q_{\star}(e^{i2\pi f_k}) = \text{sign}(c_k), \quad \forall k \in \llbracket 1, s \rrbracket \\ |Q_{\star}(e^{i2\pi f})| < 1, \quad \text{otherwise,} \end{cases} \quad (16)$$

then the solution of the Program (2) and the reconstructed estimate $\hat{x}_{\mathcal{B}}$ obtained by solving (15) are unique and verify $\hat{x}_0 = \hat{x}_{\mathcal{B}}$.

This proposition is an immediate consequence of the imbrication of the dual feasible set of Program (15) in the one Program (2), the interested reader is invited to refer to [15] for the proof details.

Finding explicit sufficient conditions for the existence of such dual certificate is a difficult problem in the general case. One might expect their existence to be related to the separability condition discussed in Section I. Although explicit criterion have been provided for specific categories of matrices, the problem for arbitrary matrices $M \in \mathbb{C}^{m \times n}$ remains an active area of research.

V. APPLICATIONS

In this last section, we discuss the benefits of the novel semidefinite formulation provided in Theorem 2 for two different categories of sub-sampling patterns M . Both of those sub-sampling patterns fall into the category of selection matrices $C_{\mathcal{I}}$ introduced in Remark 3, and therefore satisfies the conditions of Theorem 2. We explicit the advantages of the reduced SDP formulation (15) in each of those settings.

A. Random sub-sampling

Random sub-sampling was introduced in the original work [3] and is characterized as follows. The observed vector y is constructed by keeping uniformly at random each of the entries of the sampling vector $x \in \mathbb{C}^n$ independently from the others. The j^{th} observation is inserted with probability p in the vector y and discarded with probability $1 - p$. Supposing that there remain m elements at the end of the process, it was proven in [3, Theorem I.1] that

$$m \geq C \max \left\{ \log^2 \frac{n}{\delta}, s \log \frac{s}{\delta} \log \frac{n}{\delta} \right\}$$

is enough to ensure with probability at least greater than $1 - \delta$ that $\hat{x}_0 = \hat{x}_{M,A}$, provided that $\Delta_{\mathbb{T}}(\hat{x}) > \frac{4}{n-1}$.

Corollary 6. *Suppose that $\Delta_{\mathbb{T}}(\hat{x}) > \frac{4}{n-1}$ and suppose that the measurements $y \in \mathbb{C}^m$ have been acquired through a random subsampling process, then the semidefinite program (15) of dimension $m + 1 = \mathcal{O}(\max \{\log^2 \frac{n}{\delta}, s \log \frac{s}{\delta} \log \frac{n}{\delta}\})$ returns the optimal line spectrum \hat{x}_0 of Problem (2) with probability at least greater than $1 - \delta$.*

The original formulation of the problem being of dimension $n + 1$, this results bring order of magnitude changes to the computational complexity of line spectral estimation problem. Indeed, the dimension $m + 1$ of SDP (15) is poly-logarithmic on the number of initial samples n .

B. Multirate sampling systems

Multirate sampling systems (MRSS) have been studied in [15], [16] as a way to estimate sparse spectra in distributed environments. Those systems are formed by a set of p uniform samplers acquiring $\{n_l\}_{1 \leq l \leq p}$ measures at potentially different delays $\{\tau_l\}_{1 \leq l \leq p}$ and sampling frequencies $\{F_l\}_{1 \leq l \leq p}$. The observation vector $y \in \mathbb{C}^m$ is obtained by merging the different outputs of all those samplers. It is shown in [15] that if the samplers obey a common alignment property on a

grid of n_{ϕ} elements and if there exists at least one sampler of index l_* for which

$$\begin{cases} \Delta_{\mathbb{T}}\left(\hat{x}\left(\frac{\cdot}{F_{l_*}}\right)\right) \geq \frac{2.52}{n_j-1} \\ n_j > 2 \times 10^3, \end{cases} \quad (17)$$

whereby $\hat{x}\left(\frac{\cdot}{F_{l_*}}\right)$ denotes the normalized spectrum for the sampling frequency F_{l_*} , then the atomic norm relaxation of the line spectral estimation problem is tight. Since $m \ll n_{\phi}$ up to logarithmic order, Program (15) is particularly efficient in this context. In addition, it is shown that MRSS provides an efficient way to recover spectra at sub-Nyquist sampling frequencies. Applying Theorem 2 in this context gives the following corollary.

Corollary 7. *Consider a MRSS verifying the conditions (17), then the semidefinite program (15) of dimension $m + 1 \ll n_{\phi}$ returns the optimal solution of line spectral estimation problem (2). Moreover, the ground truth spectrum \hat{x}_* can be reconstructed at sub-Nyquist rates up to a spectral aliasing factor modulo $F_{\phi} \sim \prod_{l=1}^p F_l$.*

REFERENCES

- [1] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Communications on Pure and Applied Mathematics*, vol. 67, no. 6, pp. 906–956, 2014.
- [2] —, "Super-resolution from noisy data," *Journal of Fourier Analysis and Applications*, vol. 19, no. 6, pp. 1229–1254, 2013.
- [3] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7465–7490, Nov. 2013.
- [4] G. Tang, "Resolution limits for atomic decompositions via Markov-Bernstein type inequalities," *2015 International Conference on Sampling Theory and Applications, SampTA 2015*, pp. 548–552, May 2015.
- [5] C. Fernandez-Granda, "Super-resolution of point sources via convex programming," *arXiv:1507.07034*, 2015.
- [6] Y. Chi and Y. Chen, "Compressive two-dimensional harmonic retrieval via atomic norm minimization," *IEEE Transactions on Signal Processing*, vol. 63, no. 4, pp. 1030–1042, Feb. 2015.
- [7] Z. Yang, L. Xie, and P. Stoica, "Vandermonde decomposition of multilevel toeplitz matrices with application to multidimensional super-resolution," *arXiv:1505.02510*, 2015.
- [8] Y. Li and Y. Chi, "Off-the-grid line spectrum denoising and estimation with multiple measurement vectors," *IEEE Transactions on Signal Processing*, vol. 64, no. 5, pp. 1257–1269, Mar. 2016.
- [9] D. Yang, G. Tang, and M. B. Wakin, "Super-resolution of complex exponentials from modulations with unknown waveforms," *arXiv:1601.03712*, 2016.
- [10] V. Duval and G. Peyré, "Exact support recovery for sparse spikes deconvolution," *Foundations of Computational Mathematics*, vol. 15, no. 5, pp. 1315–1355, 2015.
- [11] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The Convex Geometry of Linear Inverse Problems," *Foundations of Computational Mathematics*, vol. 12, no. 6, pp. 805–849, 2012.
- [12] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic norm denoising with applications to line spectral estimation," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5987–5999, Dec. 2013.
- [13] G. Tang, B. N. Bhaskar, and B. Recht, "Near minimax line spectral estimation," *IEEE Transactions on Information Theory*, vol. 61, no. 1, pp. 499–512, Jan. 2015.
- [14] C. Carathéodory, "Über den variabilitätsbereich der fourier'schen konstanten von positiven harmonischen funktionen," *Rendiconti Del Circolo Matematico di Palermo (1884-1940)*, vol. 32, no. 1, pp. 193–217, 1911.
- [15] M. Ferreira Da Costa and W. Dai, "On Super-Resolution in Multirate Sampling Systems," *arXiv:1609.03142*.
- [16] —, "Achieving Super-Resolution in Multi-Rate Sampling Systems via Efficient Semidefinite Programming," *2016 IEEE Information Theory Workshop (ITW)*, pp. 424–428, Apr. 2016.